

CLAIMS:

1. A descriptor for the representation, from a video indexing viewpoint, of motions of a camera or any kind of observer or observing device within any sequence of frames in a video scene, said motions being at least one or several of the following basic operations : fixed, panning (horizontal rotation), tracking (horizontal transverse movement, also called travelling in the film language), tilting (vertical rotation), booming (vertical transverse movement), zooming (changes of the focal length), dollying (translation along the optical axis) and rolling (rotation around the optical axis), or any combination of at least two of these operations, wherein each of said motion types, except fixed, is oriented and subdivided into two components that stand for two different directions, and represented by means of an histogram in which the values correspond to a predefined size of displacement.

2. A descriptor according to claim 1, with which each motion type, assumed to be independent, has its own speed described in an unified way by choosing a common unit to represent it.

3. A descriptor according to claim 2, with which each motion type speed is represented by a pixel-displacement value working at the half-pixel accuracy.

4. A descriptor according to claim 3, with which, in order to work with integer values, speeds are rounded to the closest half-pixel value and multiplied by 2.

5. A descriptor according to ~~any one of claims 1 and 3~~ ^{wherein}, characterized in that the description is hierarchical, by means of a representation of the motion handled at any temporal granularity.

6. A descriptor according to claim 4, ^{wherein} ~~characterized in that~~, given a temporal window of the video data $[n_0, n_0 + N]$ (N is the total number of frames of the window) and the speeds of each motion type for each frame, the number of frames $N_{\text{motion_type}}$ in which

each motion type has a significant speed is computed and the temporal presence is represented by a percentage, defined as follows :

$$T_{\text{type of motion}} = \frac{N_{\text{type of motion}}}{N}$$

the temporal presence of all the possible motions being then represented by a

- 5 MotionTypesHistogram in which the values, between 0 and 100, correspond to a percentage, the values being only 0 or 100, depending on the fact that the given movement is present or not in the frame, when the window is reduced to a single frame.

- a 7. Application of a descriptor according to ~~any one of Claims 1 to 6~~ ^{wherein} to the
10 implementation of an image retrieval system comprising a camera for the acquisition of the video sequences, a video indexing device, a database, a graphical user interface, for carrying out a requested retrieval from the database, and a video monitor for displaying the retrieved information, the indexing operation within said video indexing device being based on the categorization resulting from the use of said descriptor of camera motions.

00141570.013400